

## 생체지표를 활용한 웹기반의 실험동물 군(郡) 분리 프로그램

김창환, 이대상\*

# Web Program for Laboratory Animal Group Separation Based on Biological Characteristics

Changhwan Kim and Daesang Lee\*

접수: 2012년 2월 1일 / 게재승인: 2012년 2월 22일

© 2012 The Korean Society for Biotechnology and Bioengineering

**Abstract:** The laboratory animal group separation is dividing animal population into subgroups, which have similar average and standard deviation values among the subgroups, based on the biological characteristics such as body weight, glucose level in blood, *etc.* Although group separation is very important and initial step in experimental design, it needs a labor intensive process for researchers because of making similar average and standard deviation values among the subgroups using the raw biological characteristics. To reduce the labor cost and increase the efficiency of animal grouping, we developed a web program named as laboratory animal group separation (LAGS) program. This LAGS uses biological characteristics of population, number of group, and the number of elements per each subgroup as input data. The LAGS automatically separates the population into each subgroup that has similar statistical data such as average and standard deviation values among subgroups. It also provides researchers with the extraordinary data generated in the process of grouping and the final grouping results by graphical display. Through our LAGS, researchers can validate and confirm results of laboratory animal group separation by just a few mouse clicks.

**Keywords:** Data visualization, Group separation, Laboratory

animal, Sorting, Subgroup

### 1. 서론

새로운 약물을 개발하거나 물질의 독성을 평가하기 위해서 실험동물을 이용한 시험이 필수적으로 수행되게 된다. 이러한 시험들은 일반적으로 대조물질 투여군과 시험물질 투여군으로 나누어 군을 구성하고 시험물질 투여군은 다시 용량 단계에 따라 몇 개의 군으로 구분하기도 한다. 이러한 군분리 (Group separation)의 과정은 일정한 기준에 의하여 수행하게 되는데 일반적으로 목적하는 생체수치의 평균이 유사하고 표준편차가 최소화되도록 시험군을 구성한다. 일반적인 독성 및 효력시험에서는 대조군을 제외하고 저용량, 중용량 및 고용량군으로 시험군을 구분하고 시험물질에 의한 효력 및 독성을 용량상관적으로 비교·평가하게 된다. 적절한 군분리가 이루어지지 않았을 경우, 시험물질의 독성 혹은 약물의 효력평가에 치명적인 오류를 범할 수 있다.

군분리는 시험물질의 독성을 평가하기 위해 수행하는 독성시험에서 가장 빈번하게 이용이 된다. 독성시험이란 시험물질 (의약, 농약, 화장품, 화학물질)을 시험체 (동물 혹은 세포 등)에 적용하여 발생하는 부작용 (독성, adverse effect)을 양적·질적으로 평가하는 시험이다. 이는 약물의 개발 및 시험물질의 안전성 평가에 반드시 필요한 시험과정이며 시험물질에 따라 각 기관에서 정한 일정한 기준에 따라 시험을 수행한다 [1-3]. 일반적으로 독성시험에서는 실험동물의 체중을 기준으로 군분리를 실시하게 된다. 시험 시작체중의 구간 평균 및 표준편차의 차이가 최소화되도록 설정한 후, 일정한 주기를 두고 체중의 구간 비교를 실시함으로써

국방과학연구소 제 5기술연구본부 4부  
The 5<sup>th</sup> R&D Institute-4, Agency for Defense Development, North-Yuseong Road 488-160, Yuseong, DaeJeon 305-152, Korea  
Tel: +82-42-821-4683, Fax: +82-42-823-3400  
e-mail: dslee@add.re.kr

시험물질에 의한 부작용을 평가하게 된다. 예를 들어 독성 시험에서 부작용의 주요한 판단기준이 되는 최대내성용량 (Maximum Tolerable Dose, MTD)은 시작체중의 10% 이상이 감소하였을 경우가 포함된다 [4].

반면, 약물의 효력을 평가하기 위한 시험의 경우에는 약물 혹은 시험목적에 따라 다양한 생체지표들을 활용하여 군 분리를 실시한다. 예를 들어, 혈구조혈인자 (erythropoietin, colony stimulating factor 등)와 같은 약물의 효력시험에서 시험약물의 투여 전에 혈구수치의 군간 차이를 최소화하는 과정이 필요하다 [5]. 또한, 당뇨치료제의 효력을 확인할 경우에는 실험동물 혈액을 채취하여 혈당을 유사하게 맞춘 후, 치료제의 투여에 의한 혈당감소를 군간 비교·관찰하여 생체 효력을 평가해야 한다 [6].

이처럼 군분리는 다양한 동물실험에서 빈번하게 이용되며 필수적으로 요구되는 과정이다. 이러한 군분리의 과정은 고가의 동물시험 관련 프로그램의 일부 항목으로 추가되는 경우가 대부분이다 [7,8]. 따라서, 실험동물 군분리만을 위하여 별도의 프로그램을 이용하는 것이 매우 제한적이다. 또한, 엑셀 등의 상용프로그램을 이용하여 프로그램화하여 실시하기도 하는데 이러한 경우에도 군분리를 위한 시간이 많이 소요가 되고 여러 단계의 군분리 과정에서 실험개체를 혼동하여 오류를 범하는 경우도 자주 발생하게 된다.

생체지표를 활용한 실험동물 군분리 프로그램은 실험동물의 군분리에 가장 많이 이용되는 ‘군’자형 군분리법을 웹 환경을 통해 사용자가 손쉽게 사용할 수 있는 Laboratory Animal Group Separation (LAGS) 프로그램을 개발하였다. 개발된 프로그램을 활용하여 실험동물을 이용한 독성 및 효력시험에서 신속하고 편리하고 정확하게 군분리를 수행할 수 있을 것으로 기대된다.

## 2. 재료 및 방법

### 2.1. 개발환경

LAGS 프로그램은 개발언어로는 Perl을 사용하였으며 리눅스 기반의 운영체제하의 웹 서버가 구동되는 환경에서 운용이 가능토록 하였다 [9]. 사용자가 웹으로 접근하여 분석을 수행할 수 있는 버전의 프로그램과 단독 (stand alone)으로 설치하여 콘솔 (console)창에서 명령어 수행을 통해 분석을 진행 할 수 있는 프로그램 등 두 가지 버전의 프로그램으로 코딩을 하였다 [10].

### 2.2. 주요기능

LAGS 프로그램의 전체 시스템에서의 자료의 흐름도는 Fig. 1과 같다. 실험동물의 체중, 혈당치 등의 생체지표자료가 저장된 첨부파일을 업로드 하거나 텍스트박스를 통해 사용자가 직접 입력한 자료를 원시자료 (Raw data)로 사용 하였다. 사용자가 분류하고자 하는 그룹의 수와 각 그룹 당 개체 수를 입력받도록 시스템을 설계하였다. LAGS 프로그램의 초기 입력화면은 Fig. 2와 같다. 또한, 사용자가 입력한 그룹의 수와 그룹 당 마리수를 곱한 값이 사용자가 초기에 입력한

원시자료들의 숫자보다 작을 경우 오류 메시지를 화면에 출력하도록 구현하였다.

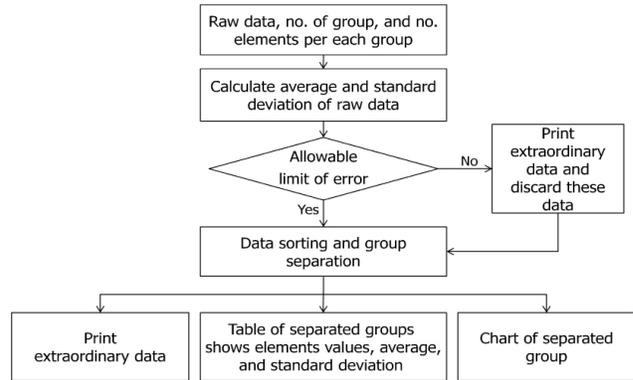


Fig. 1. Overview and schematic data flow process in the LAGS.

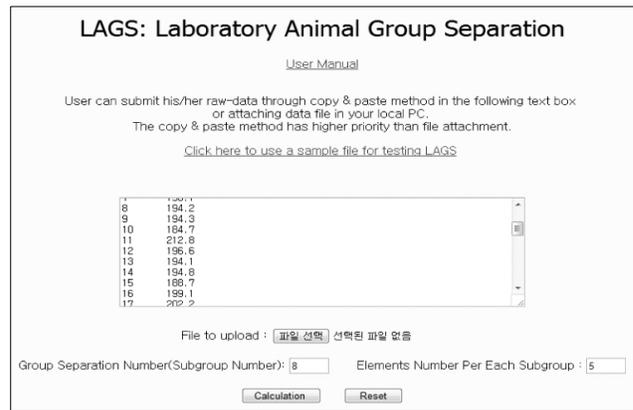


Fig. 2. Screenshot shows the main page of the LAGS.

다음 단계로는, 사용자가 직접 입력 또는 첨부한 파일의 자료로부터 읽어 들인 원시자료를 이용하여 전체 자료의 평균과 표준편차를 계산하도록 하였다. 이후, 전체 입력 자료의 표준편차에 해당하는 값 보다 2배를 초과 하거나 2배 미만 되는 값들, 즉 생체지표의 값이 평균에서 일반적으로 허용되는 오차범위를 초과하는 원시자료들의 값을 화면에 출력해 주도록 하였다. 이를 통해 사용자는 자신이 입력한 원시자료들 가운데 향후 동물실험을 수행 하는데 있어 오차 범위를 초과하는 값을 가진 실험동물을 배제 할 것인지의 여부를 결정할 수 있고, 실험에 들어갈 동물들이 생체지표 값을 기준으로 잘 선별 되었는지 확인할 수 있도록 하였다.

열 개의 실험동물집단이 있고 이들 집단에 각각 5마리씩의 동물을 할당할 경우 최소 규모로 50마리에 해당되는 실험동물의 생체지표 자료가 있어야 한다. 일반적으로 실험동물을 사육하다 보면 오차범위를 과도하게 벗어나는 생체지표를 가진 개체가 생길 수 있으므로, 실제 실험에 이용하고자 하는 동물보다 많은 수를 사육하는 경우가 대부분이다. 그 이유는 위와 같이 실험동물 50마리를 키워서 군 분리를 시도 하였는데 이들 가운데 3마리의 생체지표 값들이 표준편차의 2배 이상으로 설정한 허용 오차범위를 넘어선다면 남아있는 47마리를 가지고서는 원래 계획했던 집단으로 분리해 실험

을 수행하기 어려운 상황이 발생하기 때문이다.

민약 실험동물 50마리에 대한 생체지표를 이용하여 집단 분리에 사용할 때는 50마리의 110%에 해당되는 55마리 이상의 실험동물을 사육 한 뒤, 이들 가운데 허용 오차범위를 벗어나는 개체에 대한 자료를 배제하여 최종적으로 허용 오차범위 내에 있는 생체지표를 가진 50마리에 대한 생체 자료만을 이용하여 실험을 수행하는 것이 실험의 신뢰도를 높이는 방법이다. 따라서 LAGS 프로그램에서도 이러한 실무적인 상황을 반영하여 앞서 예를 든 55마리에 대해 최종적으로 50마리만 선별해야 하는 상황을 감안하였다. 이를 위하여, 원시자료 전체의 평균과 표준편차 값을 구한 후, 개별 개체의 생체지표들과 비교하여 표준편차가 큰 자료부터 순차적으로 배제하였고, 배제된 값들을 사용자에게 알려주도록 알고리즘을 고안하였다.

다음 단계로, 원시 자료의 생체지표 값들을 내림차순으로 정렬하고 이들을 각각의 집단으로 분류할 때, 각 집단의 평균 값과 표준편차의 값들이 비슷해 지도록 원시자료 값을 각각의 집단으로 할당하였다 (Fig. 3). 예를 들면, 원시자료들의 이름이 N1에서 N50까지 50개의 이름을 가진 원시자료가 있고 이들의 생체지표 값에 해당되는 요소 (element) 값들은 E1에서 E50이라는 값들을 가지고 있다고 가정을 하자. 이들을 10개의 집단 (G1에서 G10)으로 군분리를 한다고 가정하면 각각의 집단에는 5개씩의 원시자료의 이름과 요소 값을 할당해야 한다. 최종 목표는 이때 각각 5개의 원시자료의 요소 값을 포함하고 있는 개별 집단 G1에서 G10까지의 요소 값들의 평균과 표준편차가 가급적 비슷해지도록 해야 한다는 것이다.

[ Laboratory Animal Group Separation ]							Top	
Output file name : 7027.csv								
Count of Raw Data : 44								
Count of Data used in Group Separation : 40								
Group Separation Number (Subgroup Number): 8 Subgroups (G1-G8)								
Elements Number per each Subgroup : 5 (1-5)								
	1	2	3	4	5	MEAN	SD	
G1	212.3(21)	194.8(14)	194.5(33)	188.9(25)	188.7(15)	195.84	9.66	
G2	210.2(31)	196.6(12)	194.3(9)	191.0(4)	186.9(35)	195.80	8.84	
G3	209.9(26)	197.5(2)	194.3(38)	191.2(29)	186.2(30)	195.82	8.91	
G4	207.8(36)	197.8(37)	194.2(28)	191.3(5)	185.1(40)	195.24	8.42	
G5	204.3(6)	198.1(7)	194.2(6)	191.7(39)	184.9(44)	194.64	7.23	
G6	202.2(17)	198.1(23)	194.1(13)	192.0(34)	184.8(20)	194.24	6.57	
G7	201.2(3)	198.3(27)	193.5(32)	193.4(24)	184.7(10)	194.22	6.27	
G8	199.7(22)	199.1(16)	193.4(18)	193.4(19)	181.5(42)	193.42	7.31	

Fig. 3. The LAGS shows group separation results as table format. Each row consisted of subgroup name (G1~Gn), sorted values, element name within bracket, average and standard deviation of each subgroup.

이러한 목적을 달성하기 위해 E1에서 E50까지 무작위로 분포되어 있는 요소 값을 크기를 기준으로 내림차순 정렬하였다. 정렬과정을 통해 얻어진 값들에 대해 프로그램 내부에서 정렬 값 (Sorted value)이란 의미를 가진 S1에서 S50까지의 자료구조를 가지도록 하였다. 이러한 정렬 값들을 집단 G1에서 G10까지 할당하는 방법은 Table 1에서와 같이 실험 동물연구에 있어 일반적으로 통용되는 방식을 사용하였다.

즉, 정렬 값인 S1부터 S10까지의 값은 집단 G1에서 G10까지는 순차적으로 할당하였고, S11부터 S20까지의 값은 G10에서 G1까지 집단순서의 역방향으로 할당하였다. S21에서 S30까지의 값들은 다시 집단 G1에서 G10까지 순방향으로 할당하는 방식을 이용해 전체 집단에 정렬된 값들의 결과가 화면에 표시 되도록 하였다. 각 집단의 평균과 표준 편차를 해당 집단의 끝 부분에 별도의 칼럼을 만들어 결과 화면에 표기를 하였다. 또한 2차 자료들의 값으로부터 이들의 평균과 표준편차를 각각 계산하고 허용 오차범위에서 벗어난 값들이 있을 경우 이들을 Fig. 3에서와 같이 군 분리된 결과를 테이블 형식으로 화면에 출력하도록 개발하였다.

마지막으로, 최종 결과 화면에서 각각의 집단으로 할당된 요소 값들의 평균값과 표준편차를 막대그래프로 표시를 하였으며, 이와 함께 집단 전체의 평균의 정보와 오차 범위를 화면에 출력하여 군분리가 잘 되었는지의 유무를 사용자가 직관적으로 확인할 수 있도록 하였다 (Fig. 4). 또한, 사용자가 분석된 최종 결과 값들을 파일로 저장받아 향후 추가적으로 분석 작업이 필요할 경우 활용이 가능하도록 하였다.

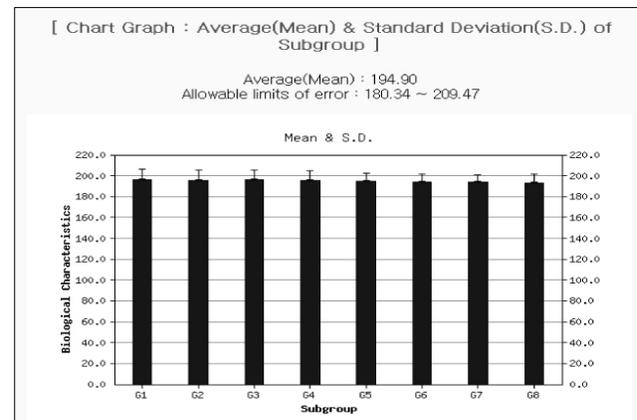


Fig. 4. The bar chart depicts the relative distribution of average among subgroups and standard deviation of subgroup is represented as error bar.

### 3. 결과 및 고찰

#### 3.1. 사례연구

LAGS 프로그램의 성능 분석을 위해 생체지표들 가운데 대표적으로 사용되고 있는 체중을 사용하여 개발시스템 (Pentium R Dual CPU 1.8 GHz, 1,000 MB RAM, Linux Fedora Core 6.0)에서 테스트 해 보았다. 이를 위하여 같은 조건에서 사육한 마흔 네마리의 6주령 수컷 SD 랫드의 체중을 원시자료로 사용하여 군분리 과정을 수행하였다. Table 2는 이들 마흔 네마리의 랫드의 체중자료이고 편의상 실험동물의 이름은 1~44까지 숫자로 표기하였다.

원시 자료 44개를 입력하고 이들을 8개의 그룹으로 나누고 각각의 그룹이 5마리씩 할당되도록 군분리를 하여 결과 값을 얻는데 소요된 시간은 개발 시스템에서 3초 내외로 나왔다. 원시 자료 44개의 1차적인 분석 결과값은 원시자료 전

체의 평균은 195.94이고, 표준 편차는 9.53으로 계산되었으며, 이들 가운데 1번, 41번, 43번 랫드가 각각 220.9, 216.8, 174.6 그램의 무게로 허용오차범위를 벗어난 값으로 나타났다. 사용자가 8개의 그룹으로 각 그룹당 5마리씩 분류되도록 사용자가 입력을 하였으므로, 군 분리에 최종적으로 필요한 자료인 40개를 초과한 4개 자료들을 배제한 결과를 도출하였다. 1차적으로 허용오차범위를 초과하는 값을 우선적으로 버리고 (1번, 41번, 43번 자료), 나머지원시 자료들 가운데 표준편차가 다른 값들에 비해 큰 값을 가지고 있는 자료 (11번 212.8 그램)를 배제하여 최종적인 군 분리 결과를 테이블과 그래프로 보여주었다. 군 분리에 최종적으로 사용된 40개의 생체지표 값들의 평균은 194.90, 표준편차는 7.28로 계산되었다.

**Table 1.** Subgroup allocation by using sorted values

Subgroup Name	Sorted values (S) and Element names (N)				
G1	S1 (N20)	S20 (N26)	S21 (N50)	S40 (N9)	S41 (N28)
G2	S2 (N31)	S19 (N4)	S22 (N27)	S39 (N38)	S42 (N3)
G3	S3 (N10)	S18 (N16)	S23 (N11)	S38 (N48)	S43 (N2)
G4	S4 (N35)	S17 (N23)	S24 (N17)	S37 (N45)	S44 (N14)
G5	S5 (N6)	S16 (N5)	S25 (N33)	S36 (N29)	S45 (N46)
G6	S6 (N15)	S15 (N12)	S26 (N49)	S35 (N36)	S46 (N47)
G7	S7 (N24)	S14 (N43)	S27 (N8)	S34 (N39)	S47 (N1)
G8	S8 (N25)	S13 (N32)	S28 (N34)	S33 (N18)	S48 (N22)
G9	S9 (N40)	S12 (N37)	S29 (N19)	S32 (N13)	S49 (N42)
G10	S10 (N7)	S11 (N44)	S30 (N41)	S31 (N21)	S50 (N30)

**3.2. 향후계획**

생체지표를 이용한 실험동물에 대한 군분리 방법은 연구실 일선에서 사용하는 전형적인 방법으로는 Excel과 같은 spreadsheet 프로그램이나 SPSS와 같은 통계 분석을 지원하는 소프트웨어를 이용하는 것이다. 군분리를 위해 Excel에 생체지표에 해당하는 원시 자료를 입력하고 서로 다른 생체지표 자료를 기준으로 평균값과 표준편차를 계산하여 허용오차범위 (표준편차의 2배 값 범위)를 벗어나는 자료를 버린다. 이후 남아 있는 자료를 이용하여 표준편차와 평균을 다시 한번 더 계산한 뒤 생체지표 자료를 정렬하여 사용자가 원하는 그룹으로 일일이 나누는 번거롭고 일손이 많이 드는 과정을 거친다. 군 분리를 위한 마지막 단계로 그룹으로 분류된 자료에 대해 검증작업하고 이를 도표화 하는 일련의 반복적이고 노동집약적인 과정을 웹으로 접속할 수 있는 환경으로 구현하여 연구자에게 편리성과 신속성을 제공할 것에 의의가 있다고 사료된다. 또한 단독 사용자 (single user) 방식의 기존 통계 분석 프로그램과는 달리 웹이라는 환경을 통해 다중 사용자 (multiuser)가 동시에 접근하여 각자가 원하는 결과를 산출할 수 있도록 하였다.

LAGS는 실험동물의 체중과 같은 생체지표를 활용하여 군 분리를 손쉽게 할 수 있도록 구현하였다. 이러한 LAGS의 특성상 본 프로그램은 중학교 또는 고등학교와 같은 교육기관의 경우 신학기 초반에 반 (class)을 편성하는데도 활용이 가능할 것으로 사료된다. 예를 들어, 생체지표에 대응되는 특

정 교과목 성적에 대해 각각의 반에서 평균과 표준편차가 비슷하게 나오도록 학생들을 배치하는 문제를 해결하는데도 LAGS가 활용될 수 있을 것으로 판단된다.

LAGS가 현재로는 원시자료의 기본적인 통계자료인 평균과 표준편차 값을 결과화면에 보여주고 있으나, 일반적인 통계 패키지를 사용할 경우에 데이터 분석 결과인 표준 오차, 분산, 최소값, 최대값, 중앙값, 최빈값, 신뢰수준 등의 수치에 대한 분석 기능 등을 향후에 LAGS에 추가할 계획이다.

**Table 2.** Animal number and its weight used in case study

Rat Number	Weight (g)	Rat Number	Weight (g)
1	220.9	23	198.1
2	197.5	24	193.4
3	202.1	25	188.9
4	191.0	26	209.9
5	191.3	27	198.3
6	204.3	28	194.2
7	198.1	29	191.2
8	194.2	30	186.2
9	194.3	31	210.2
10	184.7	32	193.5
11	212.8	33	194.5
12	196.6	34	192.0
13	194.1	35	186.9
14	194.8	36	207.8
15	188.7	37	197.8
16	199.1	38	194.3
17	202.2	39	191.7
18	193.4	40	185.1
19	193.4	41	216.8
20	184.8	42	181.5
21	212.3	43	174.6
22	199.7	44	184.9

**4. 결론**

실험동물 군분리는 동물을 이용한 생체실험에 있어 통계적 유의성 확보를 위해 실험동물이 가지고 있는 체중, 혈당과 같은 생체지표를 활용하여 동물 집단을 여러 개의 소 그룹으로 분류하고 이들 소 그룹간에 생체지표들의 평균이나 표준편차가 같게 하도록 하는 작업을 말한다. 본 논문은 실험동물을 이용한 연구를 디자인 할 때 사용자의 편의성과 군분리 결과를 직관적으로 지원할 수 있는 웹 기반의 Laboratory Animal Group Separation (LAGS) 프로그램을 개발한 것을 소개하였다. LAGS 프로그램은 웹 화면을 통해 사용자로부터 체중과 같은 생체지표에 해당되는 자료들과 이들 자료를 이용하여 분리하고자 하는 그룹의 수, 각 그룹 당 실험동물의 마리수를 입력 받는다. 분리한 각 그룹별로 생체지표 자료의 표준편차와 평균이 유사할 수 있도록 정렬한 뒤 그룹별로 분리하여 1차 및 2차의 분석결과를 테이블과 차트그래프로 도식화 하였다. 또한 군분리 과정에서 발생할 수 있는 허용범위를 초과하는 자료들에 대해 이들을 자동적으로 배제할 수 있도록 하여 최종 군분리 결과의 신뢰성을 높였다. LAGS

프로그램을 통해 연구자가 동물을 이용한 실험을 수행하는 데 있어 일상적으로 겪게 되는 군 분리 과정에서의 노동력을 줄이고 업무의 편의성을 도모하였다.

## References

1. Korea Food and Drug Administration (2009) *Guideline for administration of nonclinical test*. Notification No. 2009-183 of Korea Food and Drug Administration.
2. National Institute of Environmental Research (2009) *Guideline for the designation of research organization on the toxicity test of chemicals*. Notification No. 2009-19 of the Ministry of Environment.
3. Rural Development Administration (2009) *Guideline for the designation and operation of research organization on the safety test of agricultural chemicals*. Notification No. 2009-24 of the Rural Development Administration.
4. National Institute of Food and Drug Safety Evaluation (1999) *Standard manual the toxicity test of medicine and medical supplies*, pp. 112. Korea Food and Drug Administration, Republic of Korea.
5. Nam, C., H. Kim, J. Huh, J. Kang (1998) Efficacy of recombinant human erythropoietin. *Korean J. Lab. Anim. Sci.* 14: 241-249.
6. Dept. of Natural Drug Evaluation in KFDA (2008) *Guideline for the efficacy test of herbal medicine (oriental medicine) – diabetes mellitus*, pp. 4-5. Korea Food and Drug Administration, Republic of Korea.
7. Elvebak L. E. II. (2011) Benefits of an automated GLP final report preparation software solution. *Bioanalysis* 3: 1421-1427.
8. Kim, D., W. Son, Y. Lee, I. Kim, and N. Yang (2010) Project for the support and validation of toxicity test-Development of the computerized GLP system validation package and carcinogenicity module. *The annual report of KNTP No.* 9: 7-42.
9. Cho, Y., K. Park, and D. Lee (2008) RCA-mer: A web-based program searching for primer candidates. *Kor. J. Microbiol.* 44: 164-167.
10. Jang, J., H. Lee, M. Lee, Y. Choi, D. Choo, K. Park, and D. Lee (2009) SNPchaser: A web-based program for detecting SNPs substitution and heterozygosity existence. *KSBB Journal* 24: 410-414.